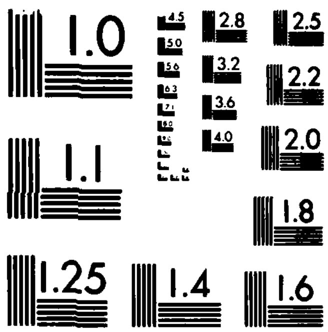
AD-A160 983 GAUSS-NEWTON METHODS FOR THE NONLINEAR COMPLEMENTARITY PROBLEM(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER P K SUBRAMANIAN AUG 85 MRC-TSR-2845 1/1

UNCLASSIFIED DAAG29-80-C-0041 F/G 12/1 NL

END


1.0
4.5
5.0
5.6
6.3
7.1
8.0
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

(2)

AD-A160 983

MRC Technical Summary Report #2845

GAUSS-NEWTON METHODS FOR THE NONLINEAR COMPLEMENTARITY PROBLEM

P. K. Subramanian

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

August 1985

(Received July 29, 1985)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, DC 20550

–A

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

GAUSS-NEWTON METHODS FOR THE NONLINEAR COMPLEMENTARITY PROBLEM

P. K. Subramanian

Technical Summary Report #2845

August 1985

ABSTRACT

It is a result of Mangasarian that the nonlinear complementarity problem is equivalent to the solution of a system of nonlinear equations. This paper considers the solution of this system by the damped Gauss-Newton method using the priniciple of least squares minimization. Algorithms are developed which under fairly simple conditions converge locally to a solution and under stronger hypotheses converge locally quadratically. The algorithms have been used successfully to solve both linear and nonlinear complementarity problems including a standard difficult test problem due to Colville.

AMS(MOS) Classification: 90C30, 90C25

Keywords: Gauss-Newton method, Complementarity problem

Work Unit Number 5: Optimization and Large Scale Systems

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based on work sponsored by National Science Foundation Grants DCR-8420963 and MCS-8102684.

SIGNIFICANCE AND EXPLANATION

Present methods of solution for the nonlinear complementarity problem consist of fixed point methods and iterative linearization. This paper proposes solution of the nonlinear complementarity problem by solving a system of nonlinear equations due to Mangasarian. Computational experience shows that the method is viable and presents an alternative approach.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

GAUSS-NEWTON METHODS

FOR THE NONLINEAR COMPLEMENTARITY PROBLEM

P. K. Subramanian

1. Introduction

Given an operator $F\colon \Re^n \longrightarrow \Re^n$, the *complementarity problem* consists in finding a z^* (if it exists) such that

$$z^* \geq 0,\ F(z^*) \geq 0 \ \text{ and } \ z^{*T}F(z^*) = 0.$$

When F is an affine operator, $F = Mx + q$ where M is an $n \times n$ matrix and $q \in \Re^n$, the complementarity problem is referred to as the *linear complementarity problem* and denoted by $LCP(M, q)$. When F is not affine, the complementarity problem is referred to simply as the *nonlinear complementarity problem* and we shall write $NLCP(F)$ in this case. The complementarity problem arises in many situations such as equilibrium problems as well as optimality conditions for mathematical programming problems (Cottle & Dantzig, [1968]).

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based on work sponsored by National Science Foundation Grants DCR-8420963 and MCS-8102684

Attempts to solve the linear complementarity problem have met with remarkable success when M is positive semidefinite or has other properties. Although the general case of $LCP(M, q)$ is NP-complete (Chung [1979]), robust methods have been found for special cases. Both *direct* and *iterative* methods are available. The best known direct methods, which are based on the process of pivoting are due to Lemke [1965] and Cottle & Dantzig [1968]. Typically, these methods end in a finite number of steps. The iterative methods are infinite methods producing a sequence of iterates whose accumulation point(s), if they exist, solve the problem. Well known amongst such mehtods are those of Cryer [1971], Cottle et al [1978] and Mangasarian [1977].

On the other hand, methods of solution for nonlinear complementarity problem are essentially iterative in nature. The best known are the *simplicial division methods* or the the so called *path following* methods due to Scarf [1967, 1973], Todd [1976] and Eaves [1972]. In such methods one essentially computes the fixed points of appropriate continuous maps such as

$$x \mapsto \left(x - F(x)\right)_+ := \min\{0, x - F(x)\}$$

where x_+ is the projection of x on the nonnegative orthant.

If F is differentiable, given the iterate x^k, we can consider the *linearization* $\mathcal{L}_k(F)$ of F at the point x^k :

$$\begin{aligned}\mathcal{L}_k(F)x &= F(x^k) + \nabla F(x^k)(x - x^k)\\ &= F(x^k) - \nabla F(x^k)x^k + \nabla F(x^k)x\end{aligned}$$

where $\nabla F(x^k)$ is the Jacobian of F at x^k. Hence if we let

$$M_k := \nabla F(x^k), \quad q_k := F(x^k) - \nabla F(x^k)x^k,$$

we get a sequence $\{x^k\}$ of iterates such that x^{k+1} solves $LCP(M_k, q_k)$. Using results from the theory of *generalized equations* developed by Robinson [1976, 1978, 1979, 1980], Josephy [1979] has shown that under suitable hypotheses, the iterates converge locally quadratically to a solution x^* of $NLCP(F)$, thus reducing $NLCP(F)$ to a sequence of linear complementarity problems. He has used this method in particular to solve Hogan's PIES model (Hogan [1975] using Lemke's algorithm to solve the resulting LCPs.

Our approach in this paper is different. In an interesting paper, Mangasarian [1976] established that $NLCP(F)$ is equivalent to the solution of a system of nonlinear equations $G(z) = 0$ for a suitably defined operator $G: \Re^n \longrightarrow \Re^n$. We shall consider several algortihms in this paper to minimize $g(z) = \|G(z)\|^2/2$. Thus our algorithms are somewhat in the spirit of least squares minimization and our methods are variants of the so called *damped Gauss-Newton procedures*. We shall show that under fairly simple conditions these algorithms exhibit local superlinear convergence and under stronger hyptheses, local quadratic convergence. Like all Newton-type methods, these methods are expensive since one needs to solve a system of linear equations at each step to find the Gauss-Newton direction. Nevertheless, used in conjunction with cheaper iterative methods to produce suitable

starting points, these methods may prove viable. We prefer to use the Gauss-Newton procedure vis-a-vis the Newton procedure since we do not require the Jacobian of G to be invertible at each iteration. In one version of our algorithm, it turns out that when ∇G is invertible our method reduces to the Newton method.

We briefly describe the notation used in this paper. We use $\Re^n$ for the space of real ordered n-tuples . All vectors are column vectors and we use the Euclidean norm throughout. Given a vector x, we denote its i^{th} component by x_i. We say $x \geq 0$ if $x_i \geq 0 \, \forall i$. The nonnegative orthant is denoted by $\Re^n_+$.

We use superscripts to distinguish between vectors, e.g., x^1, x^2 etc. For $x, y \in \Re^n$ x^T indicates the transpose of x, $x^T y$ their inner product. Occasionally, the superscript T will be suppressed. All matrices are indicated by upper case letters A, B, C etc. The i^{th} row of A is denoted by A_i while its j^{th}column is denoted by $A_{.j}$. The transpose of A is denoted by A^T.

Real valued functions defined on subsets of $\Re^n$ are denoted by lower case letters f, g, θ etc., and we write $\nabla f(x)$ and $\nabla^2 f(x)$ to indicate the gradient vector and the Hessian matrix of f at the point x. If F is an operator, $F : \Re^n \rightarrow \Re^m$, we shall write ∇F to indicate the $m \times n$ Jacobian matrix of F at a given point.

Given $NLCP(F)$, we define the *feasible set* and *solution set* by $S(F)$

and $\overline{S}(F)$ respectively, that is,

$$S(F) = \{x \in \Re^n_+ : F(x) \in \Re^n_+\}$$

$$\overline{S}(F) = \{x \in S(F) \ : \ x^T F(x) = 0\}.$$

In the case of $LCP(M, q)$, we shall denote these sets by $S(M,q)$ and $\overline{S}(M,q)$ respectively. Finally the end of a proof is signified by ▌.

2. Nonlinear equations and NLCP(F)

Our starting point for this paper is the following theorem due to Mangasarian [1976] which shows the equivalence of $NLCP(F)$ to the solution of a system of nonlinear equations.

2.1 Theorem. *Let $F: \Re^n \longrightarrow \Re^n$, θ a strictly increasing function such that $\theta: \Re \to \Re$ with $\theta(0) = 0$. Then z^* solves $NLCP(F)$ if and only if z^* solves the system of nonlinear equations $G_i(z) = 0$ $i = 1, 2 \cdots n$ where*

$$G_i(z) = \theta\big(|F_i(z) - z_i|\big) - \theta\big(F_i(z)\big) - \theta(z_i)$$

and $F_i(z)$ is the i^{th} component of $F(z)$.

We shall call $G(z)$ the *M-function* associated with F. Mangasarian also proved the following corollary which we shall find particularly useful later.

2.2 Corollary. *Suppose that z^* solves $NLCP(F)$. Assume further that F is differentiable at z^*,*

(1) *z^* is nondegenerate, that is $z^* + F(z^*) > 0$,*

(2) $\nabla F(z^*)$, *the Jacobian of F at z^* has nonsingular principal minors and*

(3) θ *is differentiable, strictly increasing such that*

$$\varsigma > 0 \Longrightarrow \theta'(\varsigma) + \theta'(0) > 0.$$

Then z^* solves $G(z) = 0$ and the Jacobian $\nabla G(z^*)$ of G is nonsingular.

3. A Gauss-Newton algorithm

Let F: $\Re^n \longrightarrow \Re^n$ and consider $NLCP(F)$. For the rest of this section we shall assume that $\nabla F(z)$ exists and the function $\theta(\varsigma)$ in Theorem 2.1 is taken to be $\theta(\varsigma) = \varsigma|\varsigma|$. Let $G(z)$ be the associated M-function. Define

$$g(z) = \frac{1}{2}\|G(z)\|^2.$$

We are going to develop an algorithm to minimize g, somewhat in the spirit of least squares minimization. We note that z^* solves $G(z) = 0$ if and only if it is a global minimizer of g. Also, if z^* is a critical point of g and $\nabla G(z^*)$ is nonsingular, then $G(z^*) = 0$ since

$$\nabla g(z^*) = \nabla G(z^*)^T G(z^*).$$

In this case, from Theorem 2.1, z^* solves $NLCP(F)$. Hence our aim is to find algorithms to find critical points of g.

Given $s \in \Re^n$, let us linearize G about s and consider (cf. Ortega & Rheinboldt, [1970], page 267)

$$g_s(x) = \frac{1}{2}\|G(s) + \nabla G(s)(x - s)\|^2.$$

Then the gradient $\nabla g_s(x)$ with respect to x of g_s is given by

$$\nabla g_s(x) = \nabla G(s)^T\big(G(s) + \nabla G(s)(x - s)\big).$$

Hence the Hessian $H_s(x)$ of g_s is given by

$$H_s(s) = \nabla G(s)^T \nabla G(s) =: A_s \qquad (3.1).$$

We note that A_s is positive semidefinite and symmetric.

3.2 Lemma, *Let* $\lambda > 0$. *For any* $x \in \Re^n$, *let* $A_x = \nabla G(x)^T \nabla G(x)$. *Suppose that* $\nabla g(x) \neq 0$. *Then the direction p given by*

$$(A_x + \lambda I)p = \nabla g(x)$$

is an ascent direction for g. *In particular,* $\exists w > 0$ *such that* $g(x - wp) < g(x)$.

Proof

There exist constants $\gamma \geq 0$ and $\nu > 0$ such that

$$\gamma\|h\|^2 \leq h^T A_x h \leq \nu\|h\|^2 \qquad \forall h \in \Re^n.$$

It follows that

$$(\gamma+\lambda)\|h\|^2 \;\leq\; h^T(A_x+\lambda I)h \;\leq\; (\nu+\lambda)\|h\|^2 \qquad \forall\, h \in \Re^n.$$

Since $\nabla g(x) \neq 0$, $p \neq 0$. If we take $h = p$, we get

$$p^T \nabla g(x) \;\geq\; (\gamma+\lambda)\|p\|^2 \;>\; 0.$$

It follows that $\nabla g(x) \cdot p \;>\; 0$ and hence [Ortega and Rheinboldt, 1970, 8.2.1] that p is an ascent direction for g. ▮

3.3 Damped Gauss-Newton Algorithm

The Lemma just proved leads us to consider the following algorithm.

(1) Let x^0 be given. Having x^k, determine x^{k+1} as follows :

(2) If $\nabla g(x^k) = 0$, stop.

(3) If $\nabla g(x^k) \neq 0$, define

$$A_k = \nabla G(x^k)^T \nabla G(x^k),$$

$$\lambda_k = g(x^k) \quad \text{and}$$

$$p^k = (A_k + \lambda_k I)^{-1} \nabla g(x^k).$$

Let ω_k be the largest element in the set $\Omega = \{1,\ 1/2, \cdots 1/2^n, \cdots\}$ such that $g(x^k - \omega_k p^k) \;<\; g(x^k)$. Set

$$x^{k+1} = x^k - \omega_k p^k.$$

3.4 Theorem. *Let x^0 be given and let $\{x^k\}$ the sequence determined by Algorithm 3.3. Assume that*

(3.5) $\sup_k \|\nabla G(x^k)\| < \infty$ *and*

(3.6) $\nabla g(x)$ *is Lipschitzian with modulus K on $\Re^n$.*

Then either $\{x^k\}$ terminates at a stationary point of g or else every limit point of $\{x^k\}$, if it exists, is a stationary point of g.

Proof

The first assertion is obvious. Let $0 < \delta \leq 1$. Let γ_k be the smallest eigenvalue of A_k. We claim that

$$\omega_k > \frac{(\lambda_k + \gamma_k)}{K}(1 - \delta). \tag{3.7}$$

If

$$\omega_k > \frac{2(\lambda_k + \gamma_k)(1 - \delta)}{K},$$

then (3.7) holds trivially. Assume, therefore, that

$$\omega_k \leq \frac{2(\lambda_k + \gamma_k)(1 - \delta)}{K}. \tag{3.8}$$

By our assumptions on $\nabla g(x)$,

$$\|\nabla g(y) - \nabla g(z)\| \leq K\|y - z\|.$$

Hence [Ortega, 1972, p 144],

$$\left|g(x^k - \omega_k p^k) - g(x^k) + \nabla g(x^k)\omega_k p^k\right| \leq \frac{K}{2} \cdot \omega_k^2 \|p^k\|^2 .$$

Hence for $\nabla g(x^k) \neq 0$ (or equivalently, $p^k \neq 0$),

$$\begin{aligned}
&g(x^k) - g(x^k - \omega_k p^k) \\
&\geq \omega_k \cdot \nabla g(x^k) p^k - \frac{K}{2} \cdot \omega_k^2 \|p^k\|^2 \\
&= \omega_k \left\{ (1-\delta) \nabla g(x^k) p^k - \frac{K\omega_k \|p^k\|^2}{2} + \delta \nabla g(x^k) \cdot p^k \right\} \\
&= \omega_k \left\{ (1-\delta) \|p^k\|^2 \left[\frac{\nabla g(x^k) p^k}{\|p^k\|^2} - \frac{K\omega_k}{2(1-\delta)} \right] + \delta \nabla g(x^k) p^k \right\}. \qquad (3.9)
\end{aligned}$$

Now by (3.5), $\exists \gamma > 0$ such that for all $h \in \Re^n$,

$$(\gamma_k + \lambda_k)\|h\|^2 \leq h^T (A_k + \lambda_k I) h \leq (\gamma + \lambda_k)\|h\|^2, \qquad (3.10a)$$

$$\frac{1}{\gamma + \lambda_k}\|h\|^2 \leq h^T (A_k + \lambda_k I)^{-1} h \leq \frac{1}{\gamma_k + \lambda_k}\|h\|^2. \qquad (3.10b)$$

Take $h = p^k$ in (3.10a) to get

$$\frac{\nabla g(x^k) p^k}{\|p^k\|^2} \geq (\lambda_k + \gamma_k).$$

Now take $h = \nabla g(x^k)$ in (3.10b) to get

$$\nabla g(x^k) p^k \geq \frac{\|\nabla g(x^k)\|^2}{(\lambda_k + \gamma)}.$$

Using (3.8) we now have that the square bracket in (3.9) is nonnegative and hence from (3.9),

$$\begin{aligned}
g(x^k) - g(x^k - \omega_k p^k) &> \delta \omega_k \nabla g(x^k) p^k \\
&\geq \frac{\delta \omega_k \|\nabla g(x^k)\|^2}{(\gamma + \lambda_k)}. \qquad (3.11)
\end{aligned}$$

Hence (3.8) implies that $g(x^k) > g(x^k - \omega_k p^k)$. Since ω_k is the largest $\omega \in \Omega$ chosen to satisfy $g(x^k) > g(x^k - \omega p^k)$, it follows that $2\omega_k$ violates (3.8) so that (3.7) holds. This proves our claim.

Assume now that x^* is a limit point of $\{x^k\}$ and that $x^{k_j} \to x^*$. From (3.11) we have

$$\begin{aligned} g(x^{k_j}) - g(x^{k_{j+1}}) &\geq g(x^{k_j}) - g(x^{k_j+1}) \\ &> \frac{\delta\omega_{k_j}\|\nabla g(x^{k_j})\|^2}{(\lambda_{k_j} + \gamma)} \\ &\longrightarrow 0. \end{aligned} \tag{3.12}$$

If $\liminf \omega_{k_j} = 0$ then by (3.7),

$$0 = \liminf \lambda_{k_j} = \liminf g(x^{k_j}) = \lim g(x^{k_j}) = g(x^*),$$

that is $G(x^*) = 0$ and hence $\nabla g(x^*) = 0$. If, however, $\liminf \omega_{k_j} = w^* > 0$, then from (3.12) and the continuity of $\nabla g(x)$, $\|g(x^*)\| = 0$, that is $\nabla g(x^*) = 0$. ∎

4. Local superlinear convergence

In this section we shall investigate situations under which conditions (3.5) and (3.6) of Theorem (3.4) can be realized. We shall then prove that under appropriate conditions, Algorithm (3.3) exhibits local superlinear convergence. We begin with the notion of a *locally Lipschitzian operator*.

4.1 Definition. *An operator $F\colon \Re^n \longrightarrow \Re^n$ is said to be locally Lipschitzian at a point z if $\exists K > 0$ and a neighborhood $N = N(z)$ of z such that F*

is Lipschitzian on N *that is*

$$\|F(x) - F(y)\| \leq K\|x - y\| \qquad \forall x,\, y \in N$$

and we write $F \in L(N)$.

4.2 Lemma. *Let* $F\colon \Re^n \longrightarrow \Re^n$ *and consider* $NLCP(F)$. *Let* $G\colon \Re^n \longrightarrow \Re^n$ *be the associated M-function,* $g(x) = \|G(x)\|^2/2$. *Suppose further that* $\nabla F_i(z)$, $i = 1, 2, \cdots n$ *is locally Lipschitzian at some point* $\bar{z}$. *Then* $\exists \delta > 0$ *such that* ∇g *is Lipschitzian on* S, *where*

$$S = S(\bar{z}, \delta) = \{x : \|x - \bar{z}\| \leq \delta\}.$$

Proof

Recall that for $i = 1, 2, \cdots n$

$$G_i(z) = \{F_i(z) - z_i\}^2 - F_i(z)|F_i(z)| - z_i|z_i| \tag{4.3}$$

so that

$$\begin{aligned}\nabla G_i(z) = 2\big(F_i(z) - z_i\big)\big(\nabla F_i(z) - e_i\big) \\ - 2|F_i(z)|\nabla F_i(z) - 2|z_i|e_i.\end{aligned} \tag{4.4}$$

Since $\nabla F_i(z) \in L(S)$, it is continuous and so are $\nabla F(z)$, $\nabla G_i(z)$ and $\nabla G(z)$. Hence all these operators are bounded on S. This implies that $F_i(z)$, $G_i(z)$ and $G(z) \in L(S)$.

Now if $A, B : D \subseteq \Re^n \rightarrow \Re^m$ and $A, B \in L(D)$, it is trivial that $A \pm B \in L(D)$. However, if A and B are also bounded on D, it follows from the identity

$$\|A(x)B(x) - A(y)B(y)\|$$
$$= \frac{1}{2}\|\big(A(x) + A(y)\big)\big(B(x) - B(y)\big) + \big(B(x) + B(y)\big)\big(A(x) - A(y)\big)\|$$

that $AB \in L(S)$. From (3.4) we now deduce that $\nabla G_i(z) \in L(S)$ and hence that $\nabla G(z) \in L(S)$. Similar considerations show that $\nabla g(z) = \nabla G(z)^T G(z) \in L(S)$. ∎

4.5 Theorem. *Let $F: \Re^n \longrightarrow \Re^n$ and consider $NLCP(F)$. Let $G: \Re^n \longrightarrow \Re^n$ be the associated M-function. Let z^* solve $NLCP(F)$, $\nabla F_i(z)$ be locally Lipschitzian at z^*, $1 \leq i \leq n$. Assume that the hypotheses of Corollary (2.2) are satisfied. Then there exists a neighborhood $N(z^*)$ such that if $x^0 \in N(z^*)$ then the sequence $\{x^k\}$ defined by Algorithm (3.3) satisfies :*

(1) $x^k \in N(z^*)$, $x^k \rightarrow z^*$,

(2) *Conditions (3.5) and (3.6) of Theorem 3.4 hold and*

(3) *If K is the constant given by (3.6), γ^* the smallest eigenvalue of $\nabla G(z^*)^T \nabla G(z^*)$ and $2\gamma^* > K$ then $x^k \rightarrow z^*$ superlinearly.*

Proof

By Corollary 2.2, $\nabla G(z^*)$ is nonsingular and by Theorem 2.1, $G(z^*) = 0$. Let $\mu_1, \mu_2, \cdots \mu_n$ be the eigenvalues of $\nabla G(z^*)$, $\Sigma = \{\mu_i : Re(\mu_i) < 0\}$

where $Re(\mu)$ denotes the real part of μ. Let

$$\eta = \begin{cases} \min\{-|\mu_i|^2/Re(\mu_i) \, : \, \mu_i \in \Sigma\} & \\ +\infty, & \text{if } \Sigma = \emptyset \, . \end{cases}$$

We can find $\delta > 0$ such that for all $z \in S = S(z^*, \delta) = \{x : \|x - z^*\| \leq \delta\}$, we have $g(z) = \|G(z)\|^2/2 < \eta$ and that $\nabla g(z)$ is Lipschitzian with modulus K on S (Lemma 4.2).

Let x^k be in S. In Algorithm 3.3, $0 < \omega_k \leq 1$ by choice and since $\lambda_k = g(x^k) < \eta$, it follows that

$$0 < \omega \leq 1, \qquad 0 \leq \frac{\lambda_k}{2 - \omega_k} < \eta.$$

It is a consequence of [Ortega and Rheinboldt, 10.2.3] that z^* is a point of attraction of $\{x^k\}$ and that $x^k \to z^*$. By Lemma 4.2 and our choice of S, it is clear that (3.5) and (3.6) hold.

Assume now that $\gamma^* > K/2$. Recall that ω_k is the largest element in Ω such that $g(x^k - \omega_k p^k) < g(x^k)$. We saw in the proof of Theorem 3.4 that

$$\omega < \frac{2(\lambda_k + \gamma_k)}{K} \Longrightarrow g(x^k - \omega p^k) < g(x^k).$$

Since $\lambda_k \to 0$, it follows that $2(\lambda_k + \gamma_k)/K \to 2\gamma^* > 1$ and hence that $\omega_k \to 1$. Thus, $\lambda_k \to 0$, $\omega_k \to 1$ and by [Ortega and Rheinboldt, p124] that $x^k \to z^*$ superlinearly. ∎

5. Local quadratic convergence

Let x^0 be given. If we assume that $\nabla G(z)$ satisfies hypothesis stronger than

(3.6), we can modify Algorithm 3.3 so that the resulting iterates converge locally quadratically.

5.1 Modified Gauss-Newton Algorithm

Consider the following algorithm in which the *perturbation parameters* λ_k are chosen slightly differently:

(1) Let x^0 be given. Having x^k, define x^{k+1} as follows:

(2) If $\nabla g(x^k) = 0$, stop.

(3) If $\nabla g(x^k) \neq 0$, let

$$A_k = \nabla G(x^k)^T \nabla G(x^k),$$

$$\gamma_k = \text{ smallest eigenvalue of } A_k.$$

(4) Define the λ_k by

$$\lambda_k = \begin{cases} 0, & \text{if } \gamma_k > 0; \\ g(x^k), & \text{otherwise .} \end{cases}$$

(5) Set $p^k = (A_k + \lambda_k I)^{-1} \nabla g(x^k)$.

(6) Let ω_k be the largest element in $\Omega = \{1, 1/2, \cdots\}$ such that $g(x^k - \omega_k p^k) < g(x^k)$.

(7) Set $x^{k+1} = x^k - \omega_k p^k$.

5.2 Theorem. *Let $F: \Re^n \longrightarrow \Re^n$ and assume that x^0 is given. Let $\{x^k\}$ be the sequence of iterates of Algorithm 5.1. Assume that :*

(5.3) $0 < \liminf_k \|\nabla G(x^k)\| \leq \sup_k \|\nabla G(x^k)\| < \infty$ *and*

(5.4) $\nabla g(x)$ *is Lipschitzian with modulus K on* $\Re^n$.

Then either $\{x^k\}$ terminates at a stationary point of g, or else every limit point of $\{x^k\}$, if there exists one is a stationary point of g.

Proof

The proof is similar to that of Theorem 3.4 and so we give only an outline. Let $0 < \delta < 1$.

Let γ_k be the smallest eigenvalue of A_k. By (5.3), there exists $\gamma > 0$ such that

$$\gamma_k \|h\|^2 \leq h^T A_k h \leq \gamma \|h\|^2$$

for all $h \in \Re^n$. As in Theorm 3.4 it is easy to show that

$$\omega_k \leq \frac{2(\lambda_k + \gamma_k)(1-\delta)}{K} \Rightarrow g(x^k - \omega_k p^k) < g(x^k),$$
$$\omega_k > \frac{(\lambda_k + \gamma_k)(1-\delta)}{K},$$
$$g(x^k) - g(x^k - \omega_k p^k) > \frac{\delta \omega_k \|\nabla g(x^k)\|^2}{(\lambda_k + \gamma)} \geq 0.$$

If x^* is a limit point of $\{x^k\}$ with $x^{k_j} \to x^*$, one then shows that

$$\omega_{k_j} \cdot \|\nabla g(x^{k_j})\| \to 0.$$

But

$$\liminf_j \omega_{k_j} > \frac{(1-\delta)}{K} \liminf_j \gamma_{k_j} > 0$$

by (5.3) so that

$$0 = \lim \|\nabla g(x^{k_j}\| = \|\nabla g(x^*)\|. \blacksquare$$

(Note that unlike Theorem 3.4, $\lambda_k = 0$ for all k is possible so that $\lambda_k \to 0 \not\Leftarrow \nabla g(x^*) = 0$. This is a consequence of our insistence that for points

sufficiently close to a nonsingular point, $A_k + \lambda_k I$ must in fact be identical to A_k, and hence $\lambda_k = 0$.)

Corresponding to Theorem 4.5 we have the following result.

5.5 Theorem. *Let F: $\Re^n \longrightarrow \Re^n$, G: $\Re^n \longrightarrow \Re^n$ the corresponding M-function. Assume that z^* solves $NLCP(F)$ and that the conditions of Corollary 2.2 hold. Assume further that $\nabla F_i(z)$, $1 \le i \le n$ is locally Lipschitzian at z^*. Then there exists a neighborhood $N(z^*)$ of z^* such that if $x^0 \in N(z^*)$ then $\{x^k\}$ defined by Algorithm 5.1 satisfies :*

(1) *$x^k \in N(z^*)$, $x^k \to z^*$ superlinearly,*

(2) *Conditions (5.3) and (5.4) of Theorem 5.4 hold and*

(3) *if K is the constant given by (5.4), γ^* the smallest eigenvalue of $\nabla G(z^*)^T \nabla G(z^*)$ and $\gamma^* > K/2$, then $x^k \to z^*$ quadratically.*

Proof

By Corollary 2.2, $\nabla G(z^*)$ is nonsingular and $G(z^*) = 0$. Using Lemma 4.2 we can find $\delta > 0$ such that for all z in $S := S(z^*, \delta) = \{x : \|x - z^*\} < \delta$, $\nabla G(z)$ is nonsingular and $\sup \|\nabla G(z)\|$ is finite.

Let $x^k \in S$, $k \ge 0$. Then A_k is positive definite so that in Algorithm 5.1, $\gamma_k > 0$ and $\lambda_k = 0$. Hence $0 < \omega_k \le 1$, $\lambda_k \equiv 0$ for all k. By [Ortega and Rheinboldt, 1970, 10.2.3], z^* is a point of attraction for $\{x^k\}$. By [Ortega and Rheinboldt, 1970, p124], $x^k \to z^*$ superlinearly. It is easy to see that (5.3) and (5.4) hold.

Suppose now that $\gamma^* > K/2$. Notice that since $\lambda_k = 0$, we have

$$\begin{aligned} x^{k+1} &= x^k - \omega_k A_k^{-1} \nabla g(x^k) \\ &= x^k - \omega_k \nabla G(x^k)^{-1} \left[\nabla G(x^k)^T\right]^{-1} \nabla G(x^k)^T G(x^k) \\ &= x^k - \omega_k \nabla G(x^k)^{-1} G(x^k). \end{aligned}$$

One now shows exactly as in the proof of Theorm 4.5 that $2(\lambda_k + \gamma_k)/K \to 2\gamma^*/K > 1$. Hence by our choice, for large k, $\omega_k = 1$. Hence Algorithm 5.1 is simply the Newton process and since $\nabla G(z)$ is Lipschitzian in S (Lemma 4.2), it follws that $x^k \to z^*$ quadratically. ▮

5.6 Remarks

1. In Algorithm 5.1, one can also take $\lambda_k = \lambda > 0$ whenever $\gamma_k = 0$ and computationally this may be preferable.

2. The robustness of the algorithm depends strongly on γ^*. If γ^* is very small, then it is possible that $\omega_k \to 0$ and the algorithm fails.

3. Although the conditions of Corollary 2.2 are only *sufficient* conditions, the algorithms developed in this paper failed to solve problems in which *none* of those conditions were fulfilled indicating, perhaps, the sharpness of those conditions. Nevertheless, many problems in which the nondegeneracy condition (Corollary 2.2, (1)) alone failed were successfully solved.

4. Algorithm 5.1 was used successfully to solve Colville test problems 1 and 2 ([Colville, 1968]). Colville problem 2 is a standard difficult test problem. The computing time in each case compared favorably with those obtained by the use of Iterative Quadratic Programming (IQP) (Garcia-

Palomares & Mangasarian, [1976]) using Lemke's algorithm to solve the quadratic subproblems : Problem 1 was solved in 2.1 seconds (0.5 seconds for IQP) and problem 2 in 5.1 seconds (4.5 seconds for IQP).

Acknowledgement. This represents a portion of the author's doctoral dissertation at the University of Wisconsin–Madison written under the supervision of Professor Olvi Mangasarian. The author is grateful to Professor Mangasarian for his continued support and encouragement.

BIBLIOGRAPHY

1. Chung, S. J. (1979). A note on the complexity of LCP. The LCP is NP-complete, University of Michigan, Ann Arbor, Department of Industrial Engineering, Technical Report No. 79-2.

2. Colville, A. R. (1968). A comparative study on nonlinear programming codes, IBM New York Scientific Report 320–2949.

3. Cottle R. W. and Dantzig, G. (1968). Complementary Pivot theory of Mahtematical Programming, *Linear Algebra and its Applications* **1**, pp 103–125.

4. Cottle, R. W., Golub, G. H. and Sacher, R. S. (1978). On the solution of large structured complementarity problems, *Applied Mathematics and Optimization* **4**, pp 347–363.

5. Cryer, C. W. (1971). The solution of a quadratic programming problem using systematic overrelaxation, *SIAM J. Control* **9**, pp 385–392.

6. Eaves, B. C. (1972). Homotopies for computation of fixed points, *Mathematical Programming* **3**, pp 1–22.

7. Garcia-Palomares, U. M. and Mangasarian, O. L. (1976). Superlinear convergent quasi-Newton algorithms for nonlinearly constrained optimization problems, *Mathematical Programming* **11**, pp 1–13.

8. Hogan, W. W. (1975). Energy Policy Models for Project Independence, *Comput. and Ops. Res.* **2**, pp 251–271.

9. Josephy, N. H. (1979). *Newton's method for generalized equations and the PIES energy model*, Ph.D. Thesis, Department of Industrial Engineering, University of Wisconsin-Madison.

10. Lemke, C. E. (1965). Bimatrix equilibrium points and mathematical programming, *Management Science* **11**, pp 681–689.

11. Mangasarian, O. L. (1976). Equivalence of the complementarity problem to a system of nonlinear equations, *SIAM J. Applied Math.* **31**, pp 89–92.

12. Mangasarian, O. L. (1977). Solution of symmetric linear complementarity problems by iterative methods, *J. Opt. Th. Applics.* **22**, pp 465–485.

13. Ortega, J. M. (1972). *Numerical Analysis, a second course*, Academic Press, New York.

14. Ortega, J. M. and Rheinboldt, W. C. (1970). *Iterative solution of nonlinear equations in several variables*, Academic Press, New York.

15. Robinson, S. M. (1976a). An Implicit-Function Theorem for Generalized Variational Inequalities, Technical Summary Report No. 1672, Mathematics Research Center, University of Wisconsin-Madison.

16. Robinson, S. M. (1976b). Stability Theory for Systems of Inequalities, Part II: Differentiable Nonlinear Systems, *SIAM J. Num. Anal.* **13**, pp 497–513.

17. Robinson, S. M. (1976d). Regularity and Stability for Convex Multivalued Functions, *Mathematics of Operations Research* **1**, pp 130–143.

18. Robinson, S. M. (1979). Generalized equations and their solutions,

Part I: Basic Theory, *Mathematical Programming Study* **10**, pp 128–141.

19. Robinson, S. M. (1980). Strongly regualr generalized equations, *Mathematics of Operations Research* **5**, pp 43–62.

20. Scarf, H. (1967). The approximation of fixed points of a continuous mapping, *SIAM J. Applied Math.* **15**, pp 1328–1343.

21. Scarf, H. (1973). *The computation of Economic Equilibria*, Yale University Press, New Haven, Connecticut.

22. Todd, M. (1976). On triangulations for computing fixed points, *Mathematical Programming* **10**, pp 322–346.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER #2845	2. GOVT ACCESSION NO. AD-A160 983
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) GAUSS-NEWTON METHODS FOR THE NONLINEAR COMPLEMENTARITY PROBLEM	5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) P. K. Subramanian	8. CONTRACT OR GRANT NUMBER(s) DCR-8420963 & MCS-8102684 DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53705	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 5 - Optimization and Large Scale Systems
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below	12. REPORT DATE August 1985
	13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, DC 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Gauss-Newton method, Complementarity problem

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

It is a result of Mangasarian that the nonlinear complementarity problem is equivalent to the solution of a system of nonlinear equations. This paper considers the solution of this system by the damped Gauss-Newton method using the principle of least squares minimization. Algorithms are developed which under fairly simple conditions converge locally to a solution and under stronger hypotheses converge locally quadratically. The algorithms have been used successfully to solve both linear and nonlinear complementarity problems including a standard difficult test problem due to Colville.

END

FILMED

12-85

DTIC